

Science as Structured Imagination

Helen De Cruz, Centre for Logic and Analytic Philosophy, University of Leuven, Kardinaal Mercierplein 2, 3000 Leuven, Belgium, Helen.DeCruz@hiw.kuleuven.be

Johan De Smedt, Department of Philosophy and Ethics, Ghent University, Blandijnberg 2, 9000 Ghent, Belgium, johan.desmedt@ugent.be

This is the penultimate version of the paper (after peer review), published in the *Journal of Creative Behavior*, vol. 44 (1), 29-44.

Abstract

This paper offers an analysis of scientific creativity based on theoretical models and experimental results of the cognitive sciences. Its core idea is that scientific creativity—like other forms of creativity—is structured and constrained by prior ontological expectations. Analogies provide scientists with a powerful epistemic tool to overcome these constraints. While current research on analogies in scientific understanding focuses on near analogies, where target and source domain are close, we argue that distant analogies—where target and source domain differ widely—are especially useful in periods of intense conceptual change. To argue this point, we discuss three case studies from the history of science: early physiologists like Harvey, early evolutionary biologists like Darwin, and recent theorists on the evolution of the human mind like Mithen.

1. INTRODUCTION

What mechanisms underlie scientific creativity; what enables scientists to make significant contributions to their disciplines? The quest by philosophers of science for some rationale behind scientific discovery and creativity has been recently joined by cognitive scientists (e.g., Simonton, 2003). Both examine what guides the scientific process and in what ways it resembles or differs from ordinary, everyday thought. Experimental psychological studies suggest that creativity is not unconstrained and limitless but structured by prior assumptions (Ward et al., 2002). Given that scientists are subject to the same cognitive limitations as other people, we argue that scientific creativity is likewise structured and constrained by prior expectations. Still, the occurrence of scientific innovations (on the level of the individual scientist) or of paradigm shifts (on the level of the scientific community) clearly indicates that scientists are able to overcome these constraining factors.

Whereas the focus of recent research on scientific creativity has been on near analogies (e.g., Dunbar, 1997), we show that distant analogies play a role in the scientific creative process, especially in periods of intense conceptual change. In our view, distant analogies constitute epistemic actions, which render problems more tractable by replacing the unfamiliar conceptual space of the target domain by a more familiar and therefore more congenial source domain. We start out with a survey of structured imagination in everyday cognition, and the role of intuitive ontologies. We then examine how people overcome the constraining effects of intuitive assumptions by distant analogies that apply the structure of one domain onto a different target domain. Finally, we provide three

examples from the history of science to illustrate that analogies from widely diverging domains play an important role in scientific creativity.

2. STRUCTURED IMAGINATION

2.1 Creativity Is Structured

In his seminal study, Ward (1994) asked college students to imagine extraterrestrial animals. Their creations possessed characteristic attributes of Earth animals, such as sense organs, legs and bilateral symmetry. In one follow-up study (Ward et al., 2002), subjects were asked to imagine tools that might be used by a highly intelligent species of extraterrestrials, with the following two constraints: the tools are not to be operated by power sources, and the creatures are not to have arms, legs or other appendages comparable to Earth animals. Despite these constraints, most participants relied on typical tools, such as hammers, saws and wrenches that were only slightly modified to allow the creatures to wrap them around their heads or hold them in their mouths. Subsequent interviews with the subjects revealed that a large majority heavily relied on specific examples of animals and tools to guide their creative process. The tendency to rely on existing knowledge as a guide to creativity is termed *structured imagination* (Ward, 1994). This finding has been replicated in many studies, even in children as young as five years of age (Cacciari, Levorato & Cicogna, 1997).

Creativity in more natural settings displays the same pattern of structured imagination. Religious ideas do not exhibit an unlimited cultural variability, but are constrained by prior ontological expectations. As Boyer (2001) already observed, there are no gods that only exist on Wednesdays. Religious concepts only exhibit minimal deviations from ordinary categories. Thus, gods and other supernatural agents are invariably conceptualized as having desires, emotions and intentions; they conform to a normal belief-desire psychology. What makes them exceptional is their minimal violation of category-based expectations, such as ghosts walking through walls (agents with a normal belief-desire psychology who nevertheless violate physical expectations). Barrett and Keil (1996) found that Christian believers have difficulties representing their god consistently as an omniscient, omnipresent being: they intuitively distort stories about him to fit expectations they have of normal people, such as that he can only attend to one event at the same time. The historian of technology Basalla (1988) has amply demonstrated that newly invented devices are nearly always based on existing artifacts. This is why archaeologists (e.g., O'Brien & Lyman, 2000) can study the evolution of artifacts as if they were organisms, using the same methods as paleontologists who study the origins, gradual morphological changes and extinctions of biological species. Product design works by tinkering, not by radical restructuring, which is not always advantageous: the tendency of design engineers to pattern new devices after earlier solutions often leads to suboptimal designs (Jansson, Condoor & Brock, 1993).

2.2 Intuitive Ontologies

The discussion above shows that existing conceptual spaces constrain creativity to an important extent, but it is less clear from where these conceptual spaces originate. A growing number of studies in cognitive anthropology (e.g., Boyer, 2000) and

neuroscience (e.g., Caramazza & Mahon, 2003) suggest that the way humans parse the world is not arbitrary or even solely governed by external reality, but rather that our inductive inferences rely on intuitive ontologies—a limited set of category-based evolved expectations that emerge early in development and that guide our reasoning about physical, psychological and biological phenomena. Ontology is the philosophical study of what is, i.e., what kinds of entities there are in the world, and how different categories of entities are related to each other. The question of how particular objects relate to universal properties is an ontological question (say, a particular cat and cats as a species). Human category-based inference mechanisms are ontologies in the sense that they provide a set of expectations of how specific kinds of objects will behave—they are termed intuitive because they are not the product of deliberate reflection or scientific investigation. Examples of intuitive ontologies include folk psychology (expectations of how agents will behave), folk physics (inferences of how inanimate objects move) and intuitive biology (expectations of how living things behave and develop).

Intuitive ontologies are often associated with computationally complex survival problems, including finding food, avoiding predators and handling tools. Rapid and efficient identification and reasoning about them have significant survival and reproductive advantages. New, Cosmides and Tooby (2007), for example, demonstrated that subjects are substantially faster and more accurate at detecting changes in complex scenery when animals (e.g., pigeons) were introduced or omitted compared to inanimate objects, even vehicles, which they have been trained for years to monitor for sudden life-or-death situations in traffic. As this efficiency could not be accounted for by differences in lower-level visual characteristics or expertise, the authors assumed that people might have an advantage in animal categorization by virtue of the ancestral importance of this ability, regardless of its current utility. Based on the existing literature, to date the best candidates for intuitive ontologies include animal, plant, artifact, person and body part (see Capitani et al., 2003, for an overview). Each of these domains contains a distinct set of assumptions, which means that inductive inferences made for objects belonging to one domain cannot be transferred to another. Neuropsychological studies (e.g., Farah & Rabinowitz, 2003) indeed indicate that patients with focal brain injury sometimes lose the ability to reason about animals and plants (e.g., they forget the names of plants and animals and cannot answer questions like “do eagles lay eggs”), while their knowledge of artifacts remains intact. In contrast, some patients with semantic dementia forget what artifacts are for (e.g., “is a pencil for writing or eating”), while their knowledge of animals and plants remains excellent (Capitani et al., 2003). As imagination about organisms and tools is structured by intuitive ontologies about these domains, this likely limits the scope of our creativity, as can be gleaned from the studies on imagining extraterrestrials and their toolkit. It seems highly unlikely that our evolved cognitive capacities would be widely off the mark (Quine, 1969). Nevertheless, there are indications that intuitive ontologies are built for speed rather than accuracy. In many cases, laypeople and children are satisfied with shallow accounts that are not explanations at all, as in the case of teleological explanations (e.g., eyes are there so that we can see, rain exists so that plants can grow), where the function is seen as a necessary and sufficient explanation of the structure under consideration (Lombrozo & Carey, 2006).

2.3 Analogies in Everyday Creative Thought

How do we overcome the constraints of conceptual structures while being creative? Neuropsychological research suggests that creativity does not depend on a single cognitive process or mechanism, but rather on the interaction of several cognitive processes (Vartanian & Goel, 2007). Indeed, a variety of cognitive mechanisms underlie creativity, such as analogical and metaphorical reasoning or conceptual combination. In this paper we will focus on analogical reasoning, our ability to understand new observations or concepts by mapping the structure of existing domains onto them (Gentner, 1983), like Niels Bohr's analogy between the structure of an atom (target domain) and the structure of the solar system (source domain), the well-known orbital model. Analogies can be considered as epistemic actions: they are performed to gain insight into a problem, which turned out to be impossible through an exploration of the original conceptual space. In contrast to pragmatic actions, which are performed to alter the world because of some intended physical change (e.g., driving from home to work), epistemic actions (e.g., driving around in order to explore one's new neighborhood after moving) are mainly performed to aid and augment cognitive processes (Kirsh, 1996). Analogies enhance our cognitive processes as they widen or alter the conceptual space in which problems are phrased.

3 INTUITIVE ONTOLOGIES AND SCIENTIFIC CREATIVITY

3.1 Intelligibility

Dear (2006) has argued that modern science comes in two distinguishable guises: instrumentality, or usefulness, and intelligibility, or providing accounts of how 'things really are'. These two components together foster a profound trust in modern science. Why is science efficient? Because it is true. How do we know that scientific beliefs are true? By virtue of their effective instrumental capacities. It is a circular argument, but invisibly so. Of these two guises, intelligibility plays an important part, because science has taken over the role of natural philosophy to account for natural phenomena not just in ways that are internally consistent, but that somehow seem right, make sense or feel intuitively true. De Regt and Dieks (2005) claim that understanding constitutes the main epistemic aim of science. They contrast science with a hypothetical oracle, whose predictions always turn out to be true. Although in this case empirical adequacy is ensured, we do not speak of a great scientific success since we cannot understand how these perfect predictions are brought about. Scientists clearly want more: they want to grasp how predictions are made, and develop a feeling for the consequences of theories in concrete situations. Indeed, in the history of science, approaches were rarely pursued on the basis of instrumental effectiveness alone. Children and laypeople sometimes resist scientific ideas that they find unintelligible, because their evolved cognitive architecture cannot adequately deal with them (Bloom & Weisberg, 2007). For example, children have the useful belief that unsupported objects fall downward, but this makes it difficult for them to conceptualize the world as a sphere. As a result, children create hybrid models of the world that match their intuitive understanding of physical phenomena, but that also conform to the spherical view of our planet, such as a hollow sphere or a disc-shaped earth (Vosniadou, Skopeliti & Ikospentaki, 2004).

The effects of intuitive ontologies are not restricted to children: they linger in educated adults. In the case of intuitive physics, even physics students continue to apply the wrong

but internally consistent impetus theory to predict the outcome of physical events. For example, they erroneously predict that a ball that is being swung in a circular path will continue to fly in a circular trajectory (Liu & MacIsaac, 2005). Intuitive ontologies may also play a role in scientific understanding. Cross-culturally, people hold the intuitive belief that humans are fundamentally different from all other animals (Anggoro, Waxman & Medin, 2005). Paleoanthropologists and archaeologists, who are not immune to this, indeed often invoke exceptional mechanisms for human evolution (De Cruz & De Smedt, 2007). Our intuitive ontologies lead us to perceive some accounts as more intelligible and more epistemologically satisfying than others. Given that one of the most important aims of science is to make nature intelligible, we can expect that intuitive ontologies will continue to play a role in scientific understanding.

3.2 Analogies and Scientific Creativity

Case studies from the history of science and field studies of scientists at work indicate that scientific creativity draws on the same cognitive resources as everyday creativity (Simonton, 2003). Scientists often resort to analogies. It is useful at this point to draw a distinction between near and distant analogies. A near analogy is one in which target and source come from the same or a closely related domain. In a distant analogy, target and source come from widely diverging domains. Early work on scientific creativity tended to focus on distant analogies, such as the snake analogy mentioned by Kekulé in his discovery of the benzene ring, or Rutherford's comparison between the structure of our solar system and that of an atom (the planetary model). Yet, upon closer scrutiny, it turned out that these analogies were not crucial in these discoveries. A fundamental difficulty of the research on historical scientific creativity is that one has to rely on retrospective accounts, written years or even decades after the discovery was made. It is only in rare instances (such as Darwin's notebooks) that written documents provide a reflection of the creative process in progress. Despite these limitations, which apply to any examination of historical material, the use of non-superficial analogies is well established both in scientists at work and in laypeople (Dunbar & Blanchette 2001). Thus it seems plausible that written records provide an (albeit incomplete) record of analogical reasoning.

Current studies suggest that scientists mainly work within the bounds of their conceptual structures. Dunbar (1997) studied creative scientific thought in action in molecular biological labs. He found that near analogies form the most important source of creative insight, such as the analogy from a well-understood virus to a lesser-understood virus to predict how it would behave in specific circumstances. In contrast, distant analogies were rare and served explanatory purposes, rather than epistemic ones. For example, one molecular biologist likened a polymerase chain reaction to the well-known analogy that a group of monkeys equipped with typewriters, given enough time, will type a sentence from Hamlet. Dunbar (1997, p. 488) therefore concludes that "creative ideas and novel concepts arise through a series of small changes produced by a variety of cognitive mechanisms ... Conceptual change, like evolutionary change, is the result of tinkering." The central role of near analogies accords well with classical accounts of scientific discovery, which suggest that the bulk of scientific work takes place within the bounds of existing well-delineated conceptual structures (e.g., Lakatos, 1978). Scientific imagination, like everyday creativity, is structured. Most of scientific progress takes place

in mental hops, rather than leaps (to use the terminology of Ward, 1998). In other words, most scientific creativity exhibits relatively little deviation from the source domain, and preserves most of the properties. What role then do distant analogies play in scientific discovery if near analogies alone can explain scientific progress? The molecular biologists Dunbar (1997) investigated could draw on a wealth of well-understood mechanisms and observations. But this is not the case for scientists working in new domains where near analogies are unavailable. We propose that in these instances, distant analogies can and do play a vital role.

Take Kepler, who attempted to explain why planets further from the Sun moved more slowly within the then new heliocentric Copernican model. Discarding the ancient idea that planets moved fixed on heavenly spheres, he introduced the concept of *vis motrix*, a precursor of gravity. To explain this concept, he drew on an analogy with light. The source domain of optics was fruitful because optical phenomena were better understood than the solar system:

Let us suppose then [...] that motion is dispensed by the Sun in the same proportion as light. Now the ratio in which light spreading out from a centre is weakened is stated by the opticians. For the amount of light in a small circle is the same as the amount of light or of the solar ray in the great one. Hence, as it is more concentrated in the small circle, and more thinly spread in the great one, the measure of this thinning out must be sought in the actual ratio of the circles, both for light and for the moving power (Kepler, *Mysterium cosmographicum* 1596, cited in Gentner et al., 1997, p. 16).

Another example is Stanford and Iris Ovshinsky's invention of the threshold switch (a successor to the transistor) by analogy of the human nerve cell. During the early fifties, the Ovshinskys recognized that plasticity of the nerve cell's membrane plays a crucial role in the neuronal basis of human learning. Based on their observations, they created a thin film of amorphous disordered material as the analogue of the cell membrane, and used this mechanical analogue to create the threshold switch. During this research, the Ovshinskys ventured into a radically new domain of science, that of amorphous disordered materials, whereas other scientists working in the domain of semiconductors still focused exclusively on crystalline materials. The creative use of distant analogy in this scientific process is plausible, since Stanford Ovshinsky had a keen interest in the neurophysiology of mammals, artificial intelligence and cybernetics, and actively corresponded with scientists working in these fields (Hoddeson, 2007). These cases suggest that distant analogies can play a role in periods of intense conceptual change or in the invention of radically new technological devices, when scientists cannot rely on established examples to draw near analogies from.

Because analogies are epistemic actions, performed to make problems more tractable by mapping them onto existing conceptual spaces, we expect that the source domain will be more familiar than the target domain, rather than the other way around. For example, scientific knowledge on the human mind has only seriously improved during the last 50 years with the advent of the cognitive revolution, primarily driven by computer scientists, who attempted to construct a robust science of intelligent behavior and behavioral biology. The structure and functional properties of the human mind, and of animal minds in general, remain as yet rather poorly understood. In contrast, our knowledge of how

rigid objects behave (mechanics) has been expanding steadily since Antiquity. Since the human mind is less understood than mechanics, we should expect that mechanical analogies for the human mind will be more fruitful in the context of scientific discovery than vice versa (see 4.3).

4 DISTANT ANALOGIES AS A SOURCE OF CREATIVITY IN THE EARLY STAGES OF SCIENTIFIC CONCEPT FORMATION

In scientific domains where formerly intuitive ontologies played an important role, distant analogies may have been of crucial importance to move away from well-trodden paths. In this section, we examine how the mapping of widely diverging source domains into target domains can foster conceptual change by three case-studies: Harvey's mechanical and weather analogies in early modern physiology, Darwin's population biology and wedge analogies in evolutionary theory, and the use of mechanical analogies in the understanding of the evolution of the human mind by evolutionary psychologists and cognitive archaeologists.

4.1 Early Modern Physiology

A first example to illustrate how distant analogies can help us overcome intuitive ontologies is the use of mechanical analogies in early modern physiology. Experimental psychological studies (e.g., Inagaki & Hatano, 2004) show that people are intuitive vitalists: they believe that vital power is taken from food and water and enables living things to sustain themselves, grow and prevent illness. Young children often explain the functions of internal organs in purely vitalistic terms, e.g., the function of the heart is to sustain the life of its owner. Regardless of their cultural background, up to the age of five, both Australian and Japanese children prefer vitalistic over mechanistic explanations of bodily functions (Morris, Taplin & Gelman, 2000). This vitalistic stance is also found in the work of ancient physicians. Claudius Galenus of Pergamum (2nd century AD) made significant contributions to our knowledge of blood circulation by dissecting animals. He studied the movements of the heart, the action of the valves and the pulsative force of the arteries. According to Galen, there are two kinds of blood: the dark type, found in the venous system, served as nutrition of the body. The lighter type, found in the arterial system, carried blood that was abundant in vital spirits. However, he failed to describe human circulation, which involves the transit of blood from the right to the left ventricle through the lungs (Khan, Daya & Gowda, 2005).

It was only in the early 17th century that the precise dynamics of circulation were discovered. Prior to the 16th century, internal organs were still mainly understood in vitalistic terms. The Renaissance revival of ancient texts included treatises on mechanics, such as Vitruvius' *De architectura* (ca. 25 BC), which contained accounts of hydraulics and water pumping engines, next to Archimedes' seminal works on mechanics. As a result, knowledge of hydraulics expanded rapidly during the early modern period, enabling the draining of the Low Countries. This permitted cardiovascular physiologists of that time to draw from this well-understood domain to unravel circulation, at that time poorly understood. The Paduan anatomist Benedetti published a paper in 1502 on the action of the heart valves, which he likened to unidirectional sluice gates in a canal: "three valves are purposefully placed by nature like movable gates which by turns when the heart is contracted in emitting blood do not completely shut off its passage, for these

valves close inward” (cited in Novell, 1990, p. 397). Another Paduan anatomist, Aquapendente, compared the action of the venous valves to a dam or a mill sluice. One of his pupils, William Harvey, used a variety of analogies to reason about circulation. Although his idea that the heart was like a pair of water bellows (not a pump, as is popularly assumed) was not novel, his colorful use of analogies, often from the domains of mechanics or physics, enabled him to make a more precise formulation of how human blood circulated. In his lectures to the College of Physicians, for example, he likened the mechanism of an erection to the inflation of a glove, and the working of lungs and thorax to a bladder within a bellows. Interestingly, the lectures are in Latin but many of these analogies are written out in English. In one of his lecture notes he wrote “From the structure of the heart, it is clear that the blood is constantly carried through the lungs into the aorta as by two clackes [a kind of pump with one-way valves] of a water bellows to rays water” (cited in Novell, 1990, p. 379). In *De motu cordis*, first published in 1628, chapter 4, Harvey wonders why the ventricles and auricle contractions in the mammal heart are so well-adapted to each other and responds with two mechanical analogies: “Nor is this for any other reason than it is in a piece of machinery, in which, though one wheel gives motion to another, yet all the wheels seem to move simultaneously; or in that mechanical contrivance which is adapted to firearms, where, the trigger being touched, down comes the flint, strikes against the steel, elicits a spark, which falling among the powder, ignites it, when the flame extends, enters the barrel, causes the explosion, propels the ball, and the mark is attained—all of which incidents, by reason of the celerity with which they happen, seem to take place in the twinkling of an eye”.

In the preface of *De motu cordis* Harvey draws a microcosm-macrocosm analogy between the weather cycle as understood by Aristotle and the circulatory system. This analogy, Gregory (2001) argues, was more than a simple rhetoric device; it enabled Harvey to understand the difference between the two types of blood, venous and arterial. Whereas this distinction did not pose a problem for Galenic physicians who understood circulation mainly in vitalistic terms, it posed a dilemma for Harvey who had to make the constant interconversion of the two types of blood plausible. From what was known about anatomy at that time, Harvey was unable to explain this, but the Aristotelian weather cycle provided an apt analogy. According to Aristotle, qualitative and cyclical changes from water into air and air into water could occur by evaporation and condensation. Just as the sun in the macrocosm plays a causal role in this process, so does the heart convert the blood in the microcosm by pumping it through the lungs. This analogy permitted a closed system for circulation, whereas Galen’s system, where the heart consumes the blood, remained open. These analogies from mechanics and physics enabled early modern anatomists to steer away from the intuitive vitalism that dominated early anatomical research. The idea that complex biological systems can be represented in simple mechanical terms was a fundamental shift in physiology, which remains important to this day. This shows how closely related understanding and scientific creativity are: by using these mechanical analogies, Harvey came to understand puzzling features of the human circulatory system, thereby enlarging medical knowledge.

4.2 Early Evolutionary Biology

Humans across cultures believe animals to possess a species-typical immutable ‘essence’ that guides their behavior and development. From the age of three, children adopt an

essentialist stance when reasoning about animals and plants, and even social categories. When one asks a four-year-old whether a kangaroo, raised by goat foster parents, will hop or climb, the child correctly predicts that it will exhibit kangaroo-typical behavior (Ahn et al., 2001). Intriguingly, children are often more essentialist than adults. Five-year-olds believe that French babies brought up by English-speaking parents will grow up to speak French (Gelman, 2004). This bias leads to similar folk taxonomies across the world, organized on the basis of underlying species-typical essences rather than superficial appearances (Atran, 1998). The widely held belief that species have unchanging essences has long impeded evolutionary accounts that hold that species can evolve into new species. Virtually all taxonomists prior to Darwin were species essentialists and despite their expanding taxonomical knowledge, they made little conceptual progress (Stamos, 2005). Essentialism is difficult to rhyme with the view that species can evolve into different species, which may in part be responsible for the enduring appeal of creationism and intelligent design. Indeed, even children from non-religious families spontaneously come up with creationist accounts for the origin of species (Evans, 2001).

The great innovation of Darwin and Wallace was to move away from this essentialist stance by adopting population thinking, where species are not idealized classes of entities, but groups of individuals that differ in their ability to survive and reproduce. Population thinking is an analogy that applies Malthus' theory of human population dynamics to organisms in general. In brief, Malthus (in *An essay on the principle of population*, first published in 1798) reasoned that food resources increased more slowly than population growth. Generalizing this observation to organisms in general, Darwin realized that animals and plants too have greater reproductive potential than available resources allow. As a result, organisms will compete for the same resources, and hence heritable traits that are advantageous will spread in the population, because their bearers will outcompete those that do not possess them. In the introduction to the *Origin of species*, Darwin (1859, p. 5) made his use of this analogy explicit. Likewise, in his autobiography, Wallace, co-discoverer of the principle of natural selection, stated explicitly that he would never have hit upon his theory were it not for reading Malthus (Wallace, 1905, p. 240, 360). Prior to these authors, biologists did not notice populations but focused on individuals. Linnaeus and others made standard idealized descriptions of species of animals and plants, such as 'the daisy', 'the honeybee'. It seems remarkable that taxonomists should have overlooked the individual variation within species, given that they studied dozens of specimens before making a detailed description of a given species. Yet well into the 19th century, virtually all biologists were essentialists. This led pre-Darwinian taxonomists to ignore or downplay the natural variability that species exhibit. Our intuitive essentialism privileges the underlying hidden essence of an organism, not its outward appearance. The skillful use of distant analogies enabled Darwin and Wallace to overcome these intuitive ontological assumptions.

Darwin struggled to understand how this Malthusian population pressure could result in evolutionary change (Millman & Smith, 1997). At first, he found it hard to envisage how "a multiplication of little means" could bring about the great effect of evolutionary change. To get a better grasp, he used the mechanical analogy of the wedge:

there is a force like a hundred thousand wedges trying to force every kind of adapted structure into the gaps in the œconomy of nature, or rather forming gaps by thrusting out weaker ones. The final cause of all this wedgings, must be to sort out proper structure & adapt it to change (Darwin, 1838).

The wedge analogy also appears in the first edition of the *Origin of Species*: “The face of Nature may be compared to a yielding surface, with ten thousand sharp wedges packed close together and driven inwards by incessant blows, sometimes one wedge being struck, and then another with greater force” (Darwin, 1859, p. 67). Interestingly, Darwin dropped this analogy in subsequent editions; perhaps because by then he had developed a nearer analogy, namely between natural and sexual selection (in the latter, the selective pressures are performed by potential mates rather than the external environment). This editing process also reveals an interesting interplay between analogies as epistemic and as explanatory tools: whereas in the notebook (not intended for publication), Darwin used the wedge analogy in an epistemic context, he later adopted it as an explanatory device, which was subsequently dropped in favor of a near analogy.

4.3 The Evolution of the Human Mind

Once the target domain becomes better understood, distant analogies lose much of their epistemic usefulness and near analogies predominate. Indeed, an analysis of historical case-studies on multiple analogies in evolutionary biology (Shelley, 1999) reveals that near analogies, such as inferences from extant species to extinct ones (e.g., from ungulates to horned dinosaurs) are more common than distant analogies. As we have seen in the case of Darwin, the latter are largely restricted to the early stages of scientific creativity. Therefore, we expect that distant analogies in contemporary scientific practice are mainly restricted to areas of research that possess an as yet underexplored conceptual structure. One candidate for such a domain is the evolution of the human mind. Despite important advances in our understanding of how the human brain evolved, its evolutionary origins remain as yet poorly understood. Thus, we can expect that authors who propose models for this domain will resort to distant analogies. Cosmides and Tooby’s analogy of the Swiss army knife is a well-known distant analogy for human cognition that maps the artifactual onto the psychological domain:

The mind is probably more like a Swiss army knife than an all-purpose blade: competent in so many situations because it has a large number of components—bottle opener, corkscrew, knife, toothpick, scissors—each of which is well designed for solving a different problem (Cosmides & Tooby, 1994, p. 60).

According to them, there are sound evolutionary reasons to expect the human mind to be made up of several dedicated domain-specific cognitive systems as specialized systems are better at solving distinct problems than a single processor: “We have both corkscrews and cups because each solves a particular problem better than the other. It would be extremely difficult to open a bottle of wine with a cup or to drink from a cork-screw” (Cosmides & Tooby, 1994, p. 58). It remains to be seen whether these analogies were epistemic rather than merely explanatory in nature. Nevertheless, although they were aware of nearer analogies, such as multiple specialized cognitive systems in birds and

primates proposed by Sherry and Shacter (1987), they continued to draw analogies from the artifactual domain, such as cognitive modules like ‘elegant machines’ (Tooby & Cosmides, 1995).

The cognitive archaeologist Mithen’s *Prehistory of the Mind* (1996) provides a further illustration of the use of distant analogies for epistemic purposes in the study of the evolution of the human mind. Mithen develops two elegant analogies. One conceptualizes human evolution as a play, divided into different acts (reminiscent of Shakespeare’s analogy of human life as a play in *As you like it*); it is mainly explanatory and organizational in nature. As Dunbar (1997) already noted, distant analogies often serve purely explanatory purposes, and this seems to be the case here: “Six million years [the period of human evolution since the split between humans and chimpanzees] is a vast span of time. In order to begin comprehend it, to grasp its salient pattern of events, it helps to think of those events as constituting a play, the drama of our past” (Mithen, 1996, p. 17). Accordingly, the book’s chapters are organized into four acts, beginning with a dimly lit empty stage, representing the as yet unknown last common ancestor of humans and chimpanzees, and ending with a dramatic act representing the last 100,000 years, which sees the evolution of *Homo sapiens*, its expansion out of Africa, the appearance of art and, finally, the emergence of agriculture.

More interesting for our discussion is the second distant analogy Mithen develops, that of medieval church and cathedral architecture to understand how, according to him, human cognition changed from domain-specific to domain-general. This second analogy is much more crucial for Mithen’s understanding, as, by his own account (1996, p. 63) it played an important role in the development of his theory of human evolution, which briefly stated claims that specialized domains of intuitive knowledge (in this paper referred to as intuitive ontologies) merged in the course of human evolution. Mithen uses the different phases in medieval architecture, with which he is familiar, as source domain to explain features of the evolved human mind, an unfamiliar target domain.

During my summer vacations when a student I worked on the excavation of the medieval Benedictine Abbey of San Vincenzo in Molise, Italy. I supervised the investigation of a particularly complex building, known as the ‘South Church’. [...] We deduced that there had been five phases in all, spanning the first 1,000 years AD and culminating in an elaborate multistory building housing many of the precious relics of the Abbey. [...] When I look at the evidence about the modern mind provided by the psychologists, I am reminded of our work at the South Church of San Vincenzo—or indeed any modern church or cathedral (Mithen, 1996, 63).

Mithen discerns three phases in the evolution of the human mind, which closely correspond to three phases that are often distinguished in the history of individual churches or cathedrals. In the first phase, the mind is like a central nave (as in the simple one-room churches in late classical and early medieval times), without any specialized cognitive capacities. A second phase witnesses the building of multiple ‘chapels’ of specialized cognitive capacities around this nave, in close analogy to the building of chapels in Romanesque churches. These include domain-specific capacities for reasoning

about social life, artifacts and natural history. Pursuing this distant analogy, Mithen argues that these domains do not influence each other:

A critical design feature of these chapels is that their walls are thick and almost impenetrable to sound from elsewhere in the cathedral. There is no access between the chapels. In other words, knowledge about distinct behavioral domains cannot be combined together (Mithen, 1996, p. 69).

Hominids from this phase, such as *Homo ergaster*, cannot reason across domains, which would explain why they did not make specialized tools, but rather general-purpose tools such as handaxes. In other words, since the artifactual and biological domains could not communicate, these hominids could not develop specialized hunting tools. The third phase is marked by a partial demolition of the separating walls between the distinct cognitive domains, so that information from one domain can flow to others. Here, the analogy draws on the transition from Romanesque to Gothic architecture, where the thick, heavy walls between the chapels were replaced by thinner columns. The differences between the minds of the second and the third phase are analogous to those between Romanesque and the succeeding Gothic cathedrals.

In Gothic architecture sound and light emanating from different parts of the cathedral can flow freely around the building unimpeded by the thick heavy walls and low vaults one finds in Romanesque architecture. [...] Similarly, in the Phase 3 mental architecture, thoughts and knowledge generated by specialized intelligences can now flow freely around the mind [...] [T]he result is an almost limitless capacity for imagination. So we should refer to these Phase 3 minds as having a ‘cognitive fluidity’ (Mithen, 1996, p. 71).

Cognitive fluidity is exemplified in animism (endowing inanimate objects with a desire-belief psychology), totemism (merging social and biological domains by making animals ancestors to current human groups) and anthropomorphism (endowing animals with human properties). According to Mithen, this third phase started about 60,000 years ago, when we see the first material evidence for across-domain reasoning, including specialized hunting tools which reveal a cross-fertilization of natural history knowledge and technology (e.g., harpoons for specialized fish capturing), and the emergence of therianthropes in art (half-human, half-animal creatures such as the 33,000 year old ‘lion man’ from Hohlenstein Stadl, Germany) which reveals a crosstalk between the social and biological worlds. In *The prehistory of the mind* the medieval cathedral analogy seeks to unify two hitherto unrelated facts about human cognition in a single explanatory framework: the fact that earlier hominids did not possess specialized, standardized tools, and the propensity of cognitively modern humans to frequently cross ontological boundaries in their reasoning. In both cases, authors have probed the evolutionary origins of the mind using distant analogies for epistemic purposes. Interestingly, none of these models apply analogies from psychology; rather they draw from heterogeneous domains, notably artifacts and architectural history. Note that such explorations of conceptual space are not always successful: in 1976, Dawkins coined the term ‘meme’ in analogy with ‘gene’ as a means to study cultural transmission. However, current models of

cultural evolution have not taken up this notion of replicating autonomous units of culture.

5 CONCLUSION

We have argued that scientific creativity draws on the same cognitive resources as other types of creativity: existing conceptual structures constrain scientists in their creative process. As a consequence, scientific creativity mostly works with small incremental steps, rather than revolutionary leaps. An important class of conceptual structures are intuitive ontologies, which guide our thinking about physical, psychological and biological phenomena. As experimental studies and the history of science reveal, they sometimes impede scientific progress. To overcome these cognitive limitations, scientists can apply distant analogies in which the ontologies of source and target domains differ widely. By presenting problems in terms of a different ontological category (e.g., the phrasing of organic functions in mechanical rather than biological terms), scientists can overcome their intuitive assumptions (e.g., vitalism) and offer solutions that are not possible in the original conceptual space. This is especially useful in the early stages of scientific creativity, when intuitive ontologies still play an important part, and in poorly understood fields of inquiry.

Acknowledgments

Elements of this paper were presented at the First Biennial Conference of the Society for Philosophy of Science in Practice at the University of Twente, The Netherlands (August 23–25, 2007). We gratefully acknowledge the useful comments by members of the audience. We would also thank two anonymous reviewers and Thomas Ward for suggestions and comments.

References

- Ahn, W., Kalish, C., Gelman, S.A., Medin, D.L., Luhmann, C., Atran, S., Coley, J.D., & Shafto, P. (2001). Why essences are essential in the psychology of concepts. *Cognition*, 82, 59–69.
- Anggoro, F.K., Waxman, S.R., & Medin, D.L. (2005). The effects of naming practices on children's understanding of living things. In B.G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the XXVII annual conference of the Cognitive Science Society* (pp. 565–570). Mahwah: Lawrence Erlbaum.
- Atran, S. (1998). Folk biology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Sciences*, 21, 547–569.
- Barrett, J.L., & Keil, F.C. (1996). Conceptualizing a nonnatural entity: Anthropomorphism in god concepts. *Cognitive Psychology*, 31, 219–247.
- Basalla, G. (1988). *The evolution of technology*. Cambridge: Cambridge University Press.
- Bloom, P., & Weisberg, D.S. (2007). Childhood origins of adult resistance to science. *Science*, 316, 996–997.

Boyer, P. (2000). Natural epistemology or evolved metaphysics? Developmental evidence for early-developed, intuitive, category-specific, incomplete, and stubborn metaphysical presumptions. *Philosophical Psychology*, 13, 277–296.

Boyer, P. (2001). *Religion explained: The evolutionary origins of religious thought*. New York: Basic Books.

Cacciari, C., Levorato, M.C., & Cicogna, P. (1997). Imagination at work: Conceptual and linguistic creativity in children. In T.B. Ward, S.M. Smith, & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes* (pp. 179–208). Washington, D.C.: American Psychological Association.

Capitani, E., Laiacona, M., Mahon, B.Z., & Caramazza, A. (2003). What are the facts of semantic category-specific deficits? A critical review of the clinical evidence. *Cognitive Neuropsychology*, 20, 213–261.

Caramazza, A., & Mahon, B.Z. (2003). The organization of conceptual knowledge: The evidence from category-specific deficits. *Trends in Cognitive Sciences*, 7, 354–361.

Cosmides, L., & Tooby, J. (1994). Beyond intuition and instinct blindness: Toward an evolutionarily rigorous cognitive science. *Cognition*, 50, 41–77.

Darwin, C. (1838). *Notebook D: Transmutation of species*, unpublished manuscript.

Darwin, C. (1859). *On the origin of species by means of natural selection or the preservation of favoured races in the struggle for life*. London: John Murray.

Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.

Dear, P. (2006). *The intelligibility of nature. How science makes sense of the world*. Chicago: University of Chicago Press.

De Cruz, H., & De Smedt, J. (2007). The role of intuitive ontologies in scientific understanding—The case of human evolution. *Biology and Philosophy*, 22, 351–368.

De Regt, H.W., & Dieks, D. (2005). A contextual approach to scientific understanding. *Synthese*, 144, 137–170.

Dunbar, K. (1997). How scientists think: On-line creativity and conceptual change in science. In T.B. Ward, S.M. Smith, & J. Vaid (Eds.), *Creative thought. An investigation of conceptual structures and processes* (pp. 461–493). Washington D.C.: American Psychological Association.

Dunbar, K., & Blanchette, I. (2001). The in vivo/in vitro approach to cognition: The case of analogy. *Trends in Cognitive Sciences*, 5, 334–339.

Evans, M. (2001). Cognitive and contextual factors in the emergence of diverse belief systems: Creation versus evolution. *Cognitive Psychology*, 42, 217–266.

Farah, M.J., & Rabinowitz, C. (2003). Genetic and environmental influences on the organization of semantic memory in the brain: Is “living things” an innate category? *Cognitive Neuropsychology*, 20, 401–408.

Gelman, S.A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences*, 8, 404–409.

Gentner, D. (1983). Structure mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155–170.

Gentner, D., Brem, S., Ferguson, R.W., Markman, A.B., Levidow, B.B., Wolff, P., & Forbus, K.D. (1997). Analogical reasoning and conceptual change: A case-study of Johannes Kepler. *Journal of the Learning Sciences*, 6, 3–40.

Gregory, A. (2001). Harvey, Aristotle and the weather cycle. *Studies in History and Philosophy of Biology & Biomedical Sciences*, 32, 153–168.

Hoddeson, L. (2007). Analogy and cognitive style in the history of invention: Inventor independence and closeness of compared domains. In S. Vosniadou, D. Kayser, & A. Protopapas (Eds.), *Proceedings of the European cognitive science conference* (pp. 413–417). Mahwah: Lawrence Erlbaum.

Inagaki, K., & Hatano, G. (2004). Vitalistic causality in young children’s naive biology. *Trends in Cognitive Sciences*, 8, 356–362.

Jansson, D.G., Condoor, S.S., & Brock, H.R. (1993). Cognition in design: Viewing the hidden side of the design process. *Environment and Planning B*, 20, 257–271.

Khan, I.A., Daya, S.K., & Gowda, R.M. (2005). Evolution of the theory of circulation. *International Journal of Cardiology*, 98, 519–521.

Kirsh, D. (1996). Adapting the environment instead of oneself. *Adaptive Behavior*, 4, 415–452.

Lakatos, I. (1978). *The methodology of scientific research programmes*. *Philosophical papers* (J. Worrall & G. Currie, Eds.). Cambridge: Cambridge University Press.

Liu, X., & MacIsaac, D. (2005). An investigation of factors affecting the degree of naïve impetus theory application. *Journal of Science Education and Technology*, 14, 101–116.

- Lombrozo, T., & Carey, S. (2006). Functional explanation and the function of explanation. *Cognition*, 99, 167–204.
- Millman, A.B., & Smith, C.L. (1997). Darwin's use of analogical reasoning in theory construction. *Metaphor and Symbol*, 12, 159–187.
- Mithen, S. (1996). *The prehistory of the mind: A search for the origins of art, religion and science*. London: Thames & Hudson.
- Morris, S.C., Taplin, J.E., & Gelman, S.A. (2000). Vitalism in naive biological thinking. *Developmental Psychology*, 36, 582–595.
- New, J., Cosmides, L., & Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National Academy of Sciences of the USA*, 104, 16598–16603.
- Novell, J.R. (1990). From Da Vinci to Harvey: The development of mechanical analogy in medicine from 1500 to 1650. *Journal of the Royal Society of Medicine*, 83, 396–398.
- O'Brien, M.J., & Lyman, R.L. (2000). Darwinian evolutionism is applicable to historical archaeology. *International Journal of Historical Archaeology*, 4, 71–112.
- Quine, W.V.O. (1969). *Ontological relativity and other essays*. New York: Columbia University Press.
- Shelley, C. (1999). Multiple analogies in evolutionary biology. *Studies in History and Philosophy of Science C*, 30, 143–180.
- Sherry, D.F., & Schacter, D.L. (1987). The evolution of multiple memory systems. *Psychological Review*, 94, 439–454.
- Simonton, D.K. (2003). Scientific creativity as constrained, stochastic behavior: The integration of product, person, and process perspectives. *Psychological Bulletin*, 129, 475–494.
- Stamos, D.N. (2005). Pre-Darwinian taxonomy and essentialism. *Biology and Philosophy*, 20, 79–96.
- Tooby, J., & Cosmides, L. (1995). Foreword. In S. Baron-Cohen. *Mindblindness. An Essay on Autism and Theory of Mind* (pp. xi–xviii). Cambridge, Ma.: MIT Press.
- Vartanian, O., & Goel, G. (2007). Neural correlates of creative cognition. In C. Martindale, P. Locher, & V.M. Petrov (Eds.), *Evolutionary and neurocognitive approaches to aesthetics, creativity and the arts* (pp. 195–207). Amityville: Baywood Publishing.

Vosniadou, S., Skopeliti, I., & Ikospentaki, K. (2004). Modes of knowing and ways of reasoning in elementary astronomy. *Cognitive Development*, 19, 203–222.

Wallace, A.R. (1905). *My life: A record of events and opinions*. London: Chapman and Hall.

Ward, T.B. (1994). Structured imagination: The role of category structure in exemplar generation. *Cognitive Psychology*, 27, 1–40.

Ward, T.B. (1998). Analogical distance and purpose in creative thought: Mental leaps versus mental hops. In K. Holyoak, D. Gentner, & B. Kokinov (Eds.), *Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences* (pp. 221–230). Sofia: New Bulgarian University.

Ward, T.B., Patterson, M.J., Sifonis, C.M., Dodds, R.A., & Saunders, K.N. (2002). The role of graded category structure in imaginative thought. *Memory and Cognition*, 30, 199–216.